Constraint Satisfaction with Counting Quantifiers*

Barnaby Martin[1]**, Florent Madelaine[2] and Juraj Stacho[3]***

[1] School of Engineering and Computing Sciences, Durham University
Science Laboratories, South Road, Durham DH1 3LE, UK.
[2] Clermont Université, Université d'Auvergne,
LIMOS, BP 10448, F-63000 Clermont-Ferrand, France.
[3] DIMAP and Mathematics Institute,
University of Warwick, Coventry CV4 7AL, UK.

Abstract. We initiate the study of *constraint satisfaction problems* (CSPs) in the presence of counting quantifiers, which may be seen as variants of CSPs in the mould of *quantified CSPs* (QCSPs).

We show that a **single** counting quantifier strictly between $\exists^{\geq 1} := \exists$ and $\exists^{\geq n} := \forall$ (the domain being of size n) already affords the maximal possible complexity of QCSPs (which have **both** $\exists$ and $\forall$), being Pspace-complete for a suitably chosen template.

Next, we focus on the complexity of subsets of counting quantifiers on clique and cycle templates. For cycles we give a full trichotomy – all such problems are in L, NP-complete or Pspace-complete. For cliques we come close to a similar trichotomy, but one case remains outstanding.

Afterwards, we consider the generalisation of CSPs in which we augment the extant quantifier $\exists^{\geq 1} := \exists$ with the quantifier $\exists^{\geq j}$ ($j \neq 1$). Such a CSP is already NP-hard on non-bipartite graph templates. We explore the situation of this generalised CSP on bipartite templates, giving various conditions for both tractability and hardness – culminating in a classification theorem for general graphs.

Finally, we use counting quantifiers to solve the complexity of a concrete QCSP whose complexity was previously open.

1 Introduction

The *constraint satisfaction problem* CSP($\mathcal{B}$), much studied in artificial intelligence, is known to admit several equivalent formulations, two of the best known of which are the query evaluation of primitive positive (pp) sentences – those involving only existential quantification and conjunction – on $\mathcal{B}$, and the homomorphism problem to $\mathcal{B}$ (see, e.g., [18]). The problem CSP($\mathcal{B}$) is NP-complete in general, and a great deal of effort has been expended in classifying its complexity for certain restricted cases. Notably it is conjectured [15,6] that for all fixed $\mathcal{B}$,

* The authors gratefully acknowledge the facilities of the Fields Institute in Toronto where the some of the work on this project was done during the Thematic Program of the institute in Mathematics of Constraint Satisfaction in August 2011.

** Author supported by EPSRC grant EP/G020604/1.

*** Author supported by EPSRC grant EP/I01795X/1.

the problem CSP($\mathcal{B}$) is in P or NP-complete. While this has not been settled in general, a number of partial results are known – e.g. over structures of size at most three [25,5] and over smooth digraphs [16,1].

A popular generalisation of the CSP involves considering the query evaluation problem for *positive Horn* logic – involving only the two quantifiers, $\exists$ and $\forall$, together with conjunction. The resulting *quantified constraint satisfaction problems* QCSP($\mathcal{B}$) allow for a broader class, used in artificial intelligence to capture non-monotonic reasoning, whose complexities rise to Pspace-completeness.

In this paper, we study counting quantifiers of the form $\exists^{\geq j}$, which allow one to assert the existence of at least j elements such that the ensuing property holds. Thus on a structure $\mathcal{B}$ with domain of size n, the quantifiers $\exists^{\geq 1}$ and $\exists^{\geq n}$ are precisely $\exists$ and $\forall$, respectively. Counting quantifiers have been extensively studied in finite model theory (see [11,22]), where the focus is on supplementing the descriptive power of various logics. Of more general interest is the majority quantifier $\exists^{\geq n/2}$ (on a structure of domain size n), which sits broadly midway between $\exists$ and $\forall$. Majority quantifiers are studied across diverse fields of logic and have various practical applications, e.g. in cognitive appraisal and voting theory [10]. They have also been studied in computational complexity, e.g., in [19].

We study variants of CSP($\mathcal{B}$) in which the input sentence to be evaluated on $\mathcal{B}$ (of size $|B|$) remains positive conjunctive in its quantifier-free part, but is quantified by various counting quantifiers from some non-empty set.

For $X \subseteq \{1, \ldots, |B|\}$, $X \neq \emptyset$, the X-CSP($\mathcal{B}$) takes as input a sentence given by a conjunction of atoms quantified by quantifiers of the form $\exists^{\geq j}$ for $j \in X$. It then asks whether this sentence is true on $\mathcal{B}$. The idea to study $\{1, \ldots, |B|\}$-CSP($\mathcal{B}$) is originally due to Andrei Krokhin.

In Section 3, we consider the power of a single quantifier $\exists^{\geq j}$. We prove that for each $n \geq 3$, there is a template $\mathcal{B}_n$ of size n, such that $\exists^{\geq j}$ $(1 < j < n)$ already has the full complexity of QCSP, i.e., $\{j\}$-CSP($\mathcal{B}_n$) is Pspace-complete.

In Section 4, we go on to study the complexity of subsets of our quantifiers on clique and cycle templates, $\mathcal{K}_n$ and $\mathcal{C}_n$, respectively. We derive the following classification theorems.

Theorem 1. *For $n \in \mathbb{N}$ and $X \subseteq \{1, \ldots, n\}$:*
- *(i) X-CSP($\mathcal{K}_n$) is in L if $n \leq 2$ or $X \cap \{1, \ldots, \lfloor n/2 \rfloor\} = \emptyset$.*
- *(ii) X-CSP($\mathcal{K}_n$) is NP-complete if $n > 2$ and $X = \{1\}$.*
- *(iii) X-CSP($\mathcal{K}_n$) is Pspace-complete if $n > 2$ and either $j \in X$ for $1 < j < n/2$ or $\{1, j\} \subseteq X$ for $j \in \{\lceil n/2 \rceil, \ldots, n\}$.*

This is a near trichotomy – only the cases where n is even and we have the quantifier $\exists^{\geq n/2}$ remain open. For cycles, however, the trichotomy is complete.

Theorem 2. *For $n \geq 3$ and $X \subseteq \{1, \ldots, n\}$, the problem X-CSP($\mathcal{C}_n$) is either in L, is NP-complete or is Pspace-complete. Namely:*
- *(i) X-CSP($\mathcal{C}_n$) $\in$ L if $n = 4$, or $1 \notin X$, or n is even and $X \cap \{2, \ldots, n/2\} = \emptyset$.*
- *(ii) X-CSP($\mathcal{C}_n$) is NP-complete if n is odd and $X = \{1\}$.*
- *(iii) X-CSP($\mathcal{C}_n$) is Pspace-complete in all other cases.*

In Section 5, we consider $\{1, j\}$-CSP($\mathcal{H}$), for $j \neq 1$ on graphs. The CSP is already NP-hard on non-bipartite graph templates. We explore the situation of this generalised CSP on bipartite graph templates, giving various conditions for both tractability and hardness, using and extending results of Section 4. We are most interested here in the distinction between P and NP-hard. To understand which of these cases are Pspace-complete would include as a subclassification the Pspace-complete cases of QCSP($\mathcal{H}$), a question which has remained open for five years [21]. We give a classification theorem for graphs in fragments of the logic involving bounded use of $\exists^{\geq 2}$ followed by unbounded use of $\exists$. In the case of QCSP ($\exists^{\geq n}$ instead of $\exists^{\geq 2}$), this is perfectly natural and is explored with bounded alternations in, e.g., [8,9,17], and with bounded use of $\forall = \exists^{\geq n}$ in [7]. We prove that either there exists such a fragment in which the problem is NP-hard or for all such fragments the problem is in P.

Afterwards in Section 6, we use counting quantifiers to solve the complexity of QCSP($\mathcal{C}_4^*$), where $\mathcal{C}_4^*$ is the reflexive 4-cycle, whose complexity was previously open. Finally, in Section 7, we give some closing remarks and open problems.

2 Preliminaries

Let $\mathcal{B}$ be a finite structure over a finite signature σ whose domain B is of cardinality $|B|$. For $1 \leq j \leq |B|$, the formula $\exists^{\geq j} x\ \phi(x)$ with *counting quantifier* should be interpreted on $\mathcal{B}$ as stating that there exist at least j distinct elements $b \in B$ such that $\mathcal{B} \models \phi(b)$. Counting quantifiers generalise existential ($\exists := \exists^{\geq 1}$), universal ($\forall := \exists^{\geq |B|}$) and (weak) majority ($\exists^{\geq |B|/2}$) quantifiers. Counting quantifiers do not in general commute with themselves, viz $\exists^{\geq j} x \exists^{\geq j} y \neq \exists^{\geq j} y \exists^{\geq j} x$ (in contrast, $\exists$ and $\forall$ do commute with themselves, but not with one another).

For $\emptyset \neq X \subseteq \{1, \ldots, |B|\}$, the X-CSP($\mathcal{B}$) takes as input a sentence of the form $\Phi := Q_1 x_1 Q_2 x_2 \ldots Q_m x_m\ \phi(x_1, x_2, \ldots, x_m)$, where ϕ is a conjunction of positive atoms of σ and each Q_i is of the form $\exists^{\geq j}$ for some $j \in X$. The set of such sentences forms the logic X-pp (recall the pp is primitive positive). The yes-instances are those for which $\mathcal{B} \models \Phi$. Note that all problems X-CSP($\mathcal{B}$) are trivially in Pspace, by cycling through all possible evaluations for the variables.

The problem $\{1\}$-CSP($\mathcal{B}$) is better-known as just CSP($\mathcal{B}$), and $\{1, |B|\}$-CSP($\mathcal{B}$) is better-known as QCSP($\mathcal{B}$). We will consider also the logic $[2^m 1^*]$-pp and restricted problem $[2^m 1^*]$-CSP($\mathcal{B}$), in which the input $\{1, 2\}$-pp sentence has prefix consisting of no more than m $\exists^{\geq 2}$ quantifiers followed by any number of $\exists$ quantifiers (and nothing else).

A homomorphism from $\mathcal{A}$ to $\mathcal{B}$, both σ-structures, is a function $h : A \to B$ such that $(a_1, \ldots, a_r) \in R^{\mathcal{A}}$ implies $(h(a_1), \ldots, h(a_r)) \in R^{\mathcal{B}}$, for all relations R of σ. A frequent role will be played by the *retraction* problem Ret($\mathcal{B}$) in which one is given a structure $\mathcal{A}$ containing $\mathcal{B}$, and one is asked if there is a homomorphism from $\mathcal{A}$ to $\mathcal{A}$ that is the identity on $\mathcal{B}$. It is well-known that retraction problems are special instances of CSPs in which the constants of the template are all named [12].

In line with convention we consider the notion of hardness reduction in proofs to be polynomial many-to-one (though logspace is sufficient for our results).

2.1 Game characterisation

There is a simple game characterisation for the truth of sentences of the logic X-pp on a structure $\mathcal{B}$. Given a sentence Ψ of X-pp, and a structure $\mathcal{B}$, we define the following game $\mathscr{G}(\Psi,\mathcal{B})$. Let $\Psi := Q_1x_1Q_2x_2\ldots Q_mx_m\ \psi(x_1,x_2,\ldots,x_m)$. Working from the outside in, coming to a quantified variable $\exists^{\geq j}x$, the *Prover* (female) picks a subset B_x of j elements of B as witnesses for x, and an *Adversary* (male) chooses one of these, say b_x, to be the value of x. Prover wins iff $\mathcal{B} \models \psi(b_{x_1},b_{x_2},\ldots,b_{x_m})$. The following comes immediately from the definitions.

Lemma 1. *Prover has a winning strategy in the game $\mathscr{G}(\Psi,\mathcal{B})$ iff $\mathcal{B} \models \Psi$.*

We will often move seemlessly between the two characterisations of Lemma 1. One may alternatively view the game in the language of homomorphisms. There is an obvious bijection between σ-structures with domain $\{1,\ldots,m\}$ and conjunctions of positive atoms in variables $\{v_1,\ldots,v_m\}$. From a structure $\mathcal{B}$ build the conjunction $\phi_{\mathcal{B}}$ listing the tuples that hold on $\mathcal{B}$ in which element i corresponds to variable v_i. Likewise, for a conjunction of positive atoms ψ, let $\mathcal{D}_\psi$ be the structure whose relation tuples are listed by ψ, where variable v_i corresponds to element i. The relationship of $\mathcal{B}$ to $\phi_{\mathcal{B}}$ and ψ to $\mathcal{D}_\psi$ is very similar to that of *canonical query* and *canonical database* (see [18]), except there we consider the conjunctions of atoms to be existentially quantified. For example, $\mathcal{K}_3$ on domain $\{1,2,3\}$ gives rise to $\phi_{\mathcal{K}_3} := \exists v_1,v_2,v_3\ E(v_1,v_2) \wedge E(v_2,v_1) \wedge E(v_2,v_3) \wedge E(v_3,v_2) \wedge E(v_3,v_1) \wedge E(v_3,v_1)$. The Prover-Adversary game $\mathscr{G}(\Psi,\mathcal{B})$ may be seen as Prover giving j potential maps for element x in $\mathcal{D}_\psi$ (ψ is quantifier-free part of Ψ) and Adversary choosing one of them. The winning condition for Prover is now that the map given from $\mathcal{D}_\psi$ to $\mathcal{B}$ is a homomorphism.

In the case of QCSP, i.e. $\{1,|B|\}$-pp, each move of a game $\mathscr{G}(\Psi,\mathcal{B})$ is trivial for one of the players. For $\exists^{\geq 1}$ quantifiers, Prover gives a singleton set, so Adversary's choice is forced. In the case of $\exists^{\geq |B|}$ quantifiers, Prover must advance all of B. Thus, essentially, Prover alone plays $\exists^{\geq 1}$ quantifiers and Adversary alone plays $\exists^{\geq |B|}$ quantifiers.

3 Complexity of a single quantifier

In this section we consider the complexity of evaluating X-pp sentences when X is a singleton, i.e., we have at our disposal only a single quantifier.

(1) $\{1\}$-*CSP*$(\mathcal{B})$ *is in* NP *for all* $\mathcal{B}$*. For each* $n \geq 2$*, there exists a template* $\mathcal{B}_n$ *of size* n *such that* $\{1\}$-*CSP*$(\mathcal{B}_n)$ *is* NP*-complete.*
(2) $\{|B|\}$-*CSP*$(\mathcal{B})$ *is in* L *for all* $\mathcal{B}$*.*
(3) For each $n \geq 3$*, there exists a template* $\mathcal{B}_n$ *of size* n *such that* $\{j\}$-*CSP*$(\mathcal{B}_n)$ *is* Pspace*-complete for all* $1 < j < n$*.*

Proof. Parts (1) and (2) are well-known (see [23], resp. [20]). For (3), let $\mathcal{B}_{\mathrm{NAE}}$ be the Boolean structure on domain $\{0,1\}$ with a single ternary not-all-equal relation $R_{\mathrm{NAE}} := \{0,1\}^3 \setminus \{(0,0,0),(1,1,1)\}$. To show Pspace-completeness, we reduce from QCSP$(\mathcal{B}_{\mathrm{NAE}})$, the *quantified not-all-equal-3-satisfiability* (see [23]).

We distinguish two cases.

Case I: $j \leq \lfloor n/2 \rfloor$. Define $\mathcal{B}_n$ on domain $\{0, \ldots, n-1\}$ with a single unary relation U and a single ternary relation R. Set $U := \{0, \ldots, j-1\}$ and set

$$R := \{0, \ldots, n-1\}^3 \setminus \{(a,b,c) : \ a, b, c \text{ either all odd or all even}\}.$$

The even numbers will play the role of false 0 and odd numbers the role of true 1.

Case II: $j > \lfloor n/2 \rfloor$. Define $\mathcal{B}_n$ on domain $\{0, \ldots, n-1\}$ with a single unary relation U and a single ternary relation R. Set $U := \{0, \ldots, j-1\}$ and set

$$R := \{0, \ldots, n-1\}^3 \setminus \{(a,b,c) : \ a, b, c \leq n-j \text{ and either all odd or all even}\}.$$

In this case even numbers $\leq n-j$ play the role of false 0 and odd numbers $\leq n-j$ play the role of true 1. The $j-1$ numbers $n-j+1, \ldots, n-1$ are somehow universal and will always satisfy any R relation.

The reduction we use is the same for Cases I and II. We reduce QCSP$(\mathcal{B}_{\text{NAE}})$ to $\{j\}$-CSP$(\mathcal{B}_n)$. Given an input $\Psi := Q_1 x_1 Q_2 x_2 \ldots Q_m x_m \ \psi(x_1, x_2, \ldots, x_m)$ to the former (i.e. each Q_i is $\exists$ or $\forall$) we build an instance Ψ' for the latter. From the outside in, we convert quantifiers $\exists x$ to $\exists^{\geq j} x$. For quantifiers $\forall x$, we convert also to $\exists^{\geq j} x$, but we add the conjunct $U(x)$ to the quantifier-free part ψ.

We claim $\mathcal{B}_{\text{NAE}} \models \Psi$ iff $\mathcal{B}_n \models \Psi'$. For the $\exists$ variables of Ψ, we can see that any j witnesses from the domain B_n for $\exists^{\geq j}$ must include some element playing the role of either false 0 or true 1 (and the other $j-1$ may always be found somewhere). For the $\forall$ variables of Ψ, U forces us to choose both 0 and 1 among the $\exists^{\geq j}$ (and the other $j-2$ will come from $2, \ldots, j-1$). The result follows. $\square$

4 Counting quantifiers on cliques and cycles

4.1 Cliques: proof of Theorem 1

Recall that $\mathcal{K}_n$ is the complete irreflexive graph on n vertices.

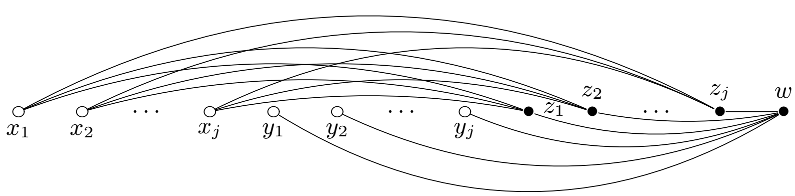

Fig. 1. The gadget $\mathcal{G}_j$.

Proposition 1. *If* $1 < j$, *then* $\{j\}$-*CSP*$(\mathcal{K}_{2j+1})$ *is* Pspace-*complete.*

Proof. By reduction from QCSP$\left(\mathcal{K}_{\binom{2j+1}{j}}\right)$, *quantified* $\binom{2j+1}{j}$-*colouring*, which is Pspace-complete by [4]. The key part of our proof involves the gadget $\mathcal{G}_j$, in Figure 1, having vertices $x_1, \ldots, x_j, y_1, \ldots, y_j, z_1, \ldots, z_j, w$ and all possible edges between $\{x_1, \ldots, x_j\}$ and $\{z_1, \ldots, z_j\}$, and between w and $\{y_1, \ldots, y_j, z_1, \ldots, z_j\}$. The left $2j$ vertices represent free variables $x_1, \ldots, x_j$, $y_1, \ldots, y_j$. Observe that $\exists^{\geq j} z_1, \ldots, z_j, w \ \ \phi_{\mathcal{G}_j}$ is true on $\mathcal{K}_{2j+1}$ iff $|\{x_1, \ldots, x_j\} \cap \{y_1, \ldots, y_j\}| < j$. If $|\{x_1, \ldots, x_j\}| = |\{y_1, \ldots, y_j\}| = j$, this is equivalent to $\{x_1 \ldots x_j\} \neq \{y_1 \ldots y_j\}$. Thus this gadget will help us to encode the edge relation on $\mathcal{K}_{\binom{2j+1}{j}}$ where we represent vertices by sets $\{a_1, \ldots, a_j\} \subset \{1, \ldots, 2j+1\}$ with $|\{a_1, \ldots, a_j\}| = j$.

Consider an instance Ψ of $\mathrm{QCSP}\big(\mathcal{K}_{\binom{2j+1}{j}}\big)$. We construct the instance Ψ' of $\{j\}$-CSP($\mathcal{K}_{2j+1}$) as follows. From the graph $\mathcal{D}_\psi$, build $\mathcal{D}_{\psi'}$ by transforming each vertex v into an independent set of j vertices $\{v^1, \ldots, v^j\}$, and each edge uv of $\mathcal{D}_\psi$ to an instance of the gadget $\mathcal{G}_j$ in which the $2j$ free variables correspond to $u^1, \ldots, u^j, v^1, \ldots, v^j$. The other variables of the gadget $\{z_1, \ldots, z_j, w\}$ are unique to each edge and are quantified innermost in Ψ' in the order $z_1, \ldots, z_j, w$.

It remains to explain the quantification of the variables of the form $v^1, \ldots, v^j$. We follow the quantifier order of Ψ. Existentially quantified variables $\exists v$ of Ψ are quantified as $\exists^{\geq j} v^1, \ldots, v^j$ in Ψ'. Universally quantified variables $\forall v$ of Ψ are also quantified $\exists^{\geq j} v^1, \ldots, v^j$ in Ψ', but we introduce additional variables $v^{1,1}, \ldots, v^{1,j+1}, \ldots, v^{j,1}, \ldots, v^{j,j+1}$ before $v^1, \ldots, v^j$ in the quantifier order of Ψ', and for each $i \in \{1, \ldots, j\}$, we join $v^{i,1}, \ldots, v^{i,j+1}$ into a clique with v^i.

It is now not difficult to verify that $\mathcal{K}_{\binom{2j+1}{j}} \models \Psi$ iff $\mathcal{K}_{2j+1} \models \Psi'$.

(For lack of space, we omit further details; please consult the appendix.) $\square$

Corollary 1. *If $1 < j < n/2$, then $\{j\}$-CSP($\mathcal{K}_n$) is* Pspace-*complete.*

Proof. We reduce from $\{j\}$-CSP($\mathcal{K}_{2j+1}$) and appeal to Proposition 1. Given an input Ψ for $\{j\}$-CSP($\mathcal{K}_{2j+1}$), we build an instance Ψ' for $\{j\}$-CSP($\mathcal{K}_n$) by adding an $(n-2j-1)$-clique on new variables, quantified outermost in Ψ', and link by an edge each variable of this clique to every other variable. Adversary chooses $n-2j-1$ elements of the domain for this clique, effectively reducing the domain size to $2j+1$ for the rest. Thus $\mathcal{K}_n \models \Psi'$ iff $\mathcal{K}_{2j+1} \models \Psi$ follows. $\square$

Proposition 2. *If $1 < j \leq n$, then $\{1, j\}$-CSP($\mathcal{K}_n$) is* Pspace-*complete.*

Proof. By reduction from QCSP($\mathcal{K}_n$). We simulate existential quantification $\exists v$ by itself, and universal quantification $\forall v$ by the introduction of $(n-j+1)$ new variables $v^1, \ldots, v^{n-j}$, joined in a clique with v, and quantified by $\exists^{\geq j}$ before v (which is also quantified by $\exists^{\geq j}$). The argument follows as in Proposition 1. $\square$

Define the n-star $\mathcal{K}_{1,n}$ to be the graph on vertices $\{0, 1, \ldots, n\}$ with edges $\{(0,j), (j,0) \ : \ j \geq 1\}$ where 0 is called the *centre* and the remainder are *leaves*.

Proposition 3. *If $X \cap \{1, \ldots, \lfloor n/2 \rfloor\} = \emptyset$, then X-CSP($\mathcal{K}_n$) is in* L.

Proof. Instance Ψ of X-CSP($\mathcal{K}_n$) of the form $\exists^{\geq \lambda_1} x_1 \ldots \exists^{\geq \lambda_m} x_m \ \psi(x_1, \ldots, x_m)$ induces the graph $\mathcal{D}_\psi$, which we may consider totally ordered (the order is given left-to-right ascending by the quantifiers). We claim that $\mathcal{K}_n \models \Psi$ iff $\mathcal{D}_\psi$ does not contain as a subgraph (not necessarily induced) a $(n-\lambda_i+1)$-star in which the $n-\lambda_i+1$ leaves all come before the centre x_i in the ordering.

($\Rightarrow$) If $\mathcal{D}_\psi$ contains such a star, then Ψ is a no-instance, as we may give a winning strategy for Adversary in the game $\mathscr{G}(\Psi, \mathcal{K}_n)$. Adversary should choose distinct values for the variables associated with the $n-\lambda_i+1$ leaves of the star (can always be done as each of the possible quantifiers assert existence of $> n/2$ elements and $n-\lambda_i < n/2$), whereupon there is no possibility for Prover to choose λ_i witnesses to the variable x_i associated with the centre.

($\Leftarrow$) If $\mathcal{D}_\psi$ does not contain such a star, then we give the following winning strategy for Prover in the game $\mathscr{G}(\Psi, \mathcal{K}_n)$. Whenever a new variable comes up, its

corresponding vertex in $\mathcal{D}_\psi$ has $l < n - \lambda_i + 1$ adjacent predecessors, which were answered with $b_1, \ldots, b_l$. Prover suggests any set of size λ_i from $B \setminus \{b_1, \ldots, b_l\}$ (which always exists) and the result follows. $\square$

Proof of Theorem 1. For $n \leq 2$ see [21], and for (ii) see [16]. The remainder of (i) is proved as Proposition 3 while Corollary 1 and Proposition 2 give (iii). $\square$

4.2 Cycles: proof of Theorem 2

Recall that $\mathcal{C}_n$ denotes the irreflexive symmetric cycle on n vertices. We consider $\mathcal{C}_n$ to have vertices $\{0, 1, \ldots, n-1\}$ and edges $\big\{(i,j) \;:\; |i-j| \in \{1, n-1\}\big\}$.

In the forthcoming proof, we use the following elementary observation from additive combinatorics. Let $n \geq 2$, $j \geq 1$, and A, B be sets of integers. Define:

- $A +_n B = \big\{(a+b) \bmod n \;\big|\; a \in A,\, b \in B\big\}$
- $j \times_n A = \underbrace{A +_n \ldots +_n A}_{j \text{ times}}$

Lemma 2. *Let $n \geq 3$ and $2 \leq j < n$. Then*

$$\Big|\, j \times_n \{-1,+1\} \,\Big| = \begin{cases} j+1 & n \text{ is odd} \\ \min\big\{j+1, n/2\big\} & n \text{ is even} \end{cases}$$

$$\Big|\, n \times_n \{-1,+1\} \,\Big| = \Big|\, n \times_n \{-2,0,+2\} \,\Big| = \begin{cases} n & n \text{ is odd} \\ n/2 & n \text{ is even} \end{cases}$$

Proposition 4. *If $n \geq 3$, then X-CSP($\mathcal{C}_n$) is in* L *if $n = 4$, or $1 \notin X$, or n is even and $X \cap \{2, 3 \ldots, n/2\} = \emptyset$,*

Proof. Let Ψ be an instance of X-CSP($\mathcal{C}_n$). Recall that $\mathcal{D}_\psi$ is the graph corresponding to the quantifier-free part of Ψ. We write $x \prec y$ if x, y are vertices of $\mathcal{D}_\psi$ (i.e., variables of ψ) such that x is quantified before y in Ψ. For an edge xy of $\mathcal{D}_\psi$ where $x \prec y$, we say that x is a *predecessor* of y. Note that a vertex can have several predecessors.

The following claims restrict the yes-instances of X-CSP($\mathcal{C}_n$).

Let x be a vertex of $\mathcal{D}_\psi$ quantified in Ψ by $\exists^{\geq j}$ for some j. If $\mathcal{C}_n \models \Psi$ then
(1a) if $j \geq 3$, then x has no predecessors,
(1b) if n is even and $j > n/2$, then x is the first vertex (w.r.t. $\prec$) of some connected component of $\mathcal{D}_\psi$, and
(1c) if $n \neq 4$ and $j = 2$, then all predecessors of x except for its first predecessor (w.r.t. $\prec$) are quantified by $\exists^{\geq 1}$.

(We omit the proof for lack of space; please consult the appendix.)

Using these claims, we prove the proposition. First, we consider the case $n = 4$. We show that $\{1,2,3,4\}$-CSP($\mathcal{C}_4$) is in L. This will imply that X-CSP($\mathcal{C}_4$) is in L for every X. Observe that if $\mathcal{D}_\psi$ contains a vertex x quantified by $\exists^{\geq 3}$ or $\exists^{\geq 4}$, then by (1b) this vertex is the first in its component (if Ψ is not a trivial no-instance). Thus replacing its quantification by $\exists^{\geq 1}$ does not change the truth of Ψ. So we may assume that Ψ is an instance of $\{1,2\}$-CSP($\mathcal{C}_4$). We now claim that $\mathcal{C}_4 \models \Psi$ if and only if $\mathcal{D}_\psi$ is bipartite. Clearly, if $\mathcal{D}_\psi$ is not bipartite, it has no homomorphism to $\mathcal{C}_4$ and hence $\mathcal{C}_4 \not\models \Psi$. Conversely, assume that $\mathcal{D}_\psi$ is bipartite with bipartition (A, B). Our strategy for Prover offers the set $\{0, 2\}$ or its subsets for the vertices in A and offers $\{1, 3\}$ or its subsets for every vertex in B. It is easy to verify that this is a winning strategy for Prover. Thus $\mathcal{C}_4 \models \Psi$. The complexity now follows as checking (1b) and checking if a graph is bipartite is in L by [24].

Now, we may assume $n \neq 4$, and next we consider the case $1 \notin X$. If also $2 \notin X$, then by (1a) the graph $\mathcal{D}_\psi$ contains no edges (otherwise Ψ is a trivial no-instance). This is clearly easy to check in L. Thus $2 \in X$. We claim that if we satisfy (1a) and (1c), then $\mathcal{C}_n \models \Psi$. We provide a winning strategy for Prover. Namely, for a vertex x, if x has no predecessors, offer any set for x. If x has a unique predecessor y for which the value i was chosen, then x is quantified by $\exists^{\geq 2}$ (or $\exists$) by (1a) and we offer $\{i-1, i+1\}$ (mod n) for x . There are no other cases by (1a) and (1c). It follows that Prover always wins with this strategy. In terms of complexity, it suffices to check (1a) and (1c) which is in L.

Finally, suppose that n is even and $X \cap \{2 \ldots n/2\} = \emptyset$. Note that every vertex of $\mathcal{D}_\psi$ is either quantified by $\exists^{\geq 1}$ or by $\exists^{\geq j}$ where $j > n/2$. Thus, using (1b), unless Ψ is a trivial no-instance, we can again replace every $\exists^{\geq j}$ in Ψ by $\exists^{\geq 1}$ without changing the truth of Ψ. Hence, we may assume that Ψ is an instance of $\{1\}$-CSP($\mathcal{C}_n$). Thus, as n is even, $\mathcal{C}_n \models \Psi$ if and only if $\mathcal{D}_\psi$ is bipartite. The complexity again follows from [24]. That concludes the proof. $\square$

Proposition 5. *Let $n \geq 3$. Then X-CSP($\mathcal{C}_n$) is Pspace-complete if $n \neq 4$ and $\{1, j\} \subseteq X$: where $j \in \{2, \ldots, n\}$ if n is odd and $j \in \{2, \ldots, n/2\}$ if n is even.*

Proof. By reduction, namely a reduction from QCSP($\mathcal{C}_n$) for odd n, and from QCSP($\mathcal{K}_{n/2}$) for even n. Both problems are known to be Pspace-hard [4].

First, consider the case of odd n. Let Ψ be an instance of QCSP($\mathcal{C}_n$). In other words, Ψ is an instance of $\{1, n\}$-CSP($\mathcal{C}_n$). Clearly, $j < n$ otherwise we are done.

We modify Ψ by replacing each universally-quantified variable x of Ψ by a path. Namely, let π_x denote the pp-formula that encodes that

$$x_1^1, x_2^1, \ldots, x_{j-1}^1,\ x_1^2, x_2^2, \ldots, x_{j-1}^2, \quad \ldots \quad , x_1^n, x_2^n, \ldots, x_{j-1}^n,\ x$$

is a path in that order (all but x are new variables). We replace $\forall x$ by

$$Q_x \ = \ \exists^{\geq j} x_1^1 \ \exists^{\geq j} x_1^2 \ldots \ \exists^{\geq j} x_1^n \ \exists^{\geq j} x \ \exists^{\geq 1} x_2^1 \ldots \exists^{\geq 1} x_{j-1}^1 \ \ldots \ \exists^{\geq 1} x_2^n \ldots \exists^{\geq 1} x_{j-1}^n$$

and append π_x to the quantifier-free part of the formula. Let Ψ' denote the final formula after considering all universally quantified variables. Note that Ψ' is an instance of $\{1, j\}$-CSP($\mathcal{C}_n$). We claim that $\mathcal{C}_n \models \Psi$ if and only if $\mathcal{C}_n \models \Psi'$.

To do this, it suffices to show that Ψ' correctly simulates the universal quantifiers of Ψ. Namely, we prove that $\mathcal{C}_n \models Q_x \pi_x$, and for each $\ell \in \{0 \ldots n-1\}$, Adversary has a strategy on $Q_x \pi_x$ that evaluates x to ℓ.

(We omit further details for lack of space; please consult the appendix.)

It remains to investigate the case of even n. Recall that $n \geq 6$ and $j \leq n/2$. We show a reduction from QCSP($\mathcal{K}_{n/2}$) to $\{1, j\}$-QCSP($\mathcal{C}_n$). The reduction is a variant of the construction from [13] for the problem of retraction to even cycles.

Let Ψ be an instance of QCSP($\mathcal{K}_{n/2}$), and define $r = (-n/2 - 2) \bmod (j-1)$. We construct a formula Ψ' from Ψ as follows. First, we modify Ψ by replacing universal quantifiers exactly as in the case of odd n. Namely, we define Q_x and π_x as before, replace each $\forall x$ by Q_x, and append π_x to the quantifier-free part of the formula. After this, we append to the formula a cycle on n vertices $v_0, v_1, \ldots, v_{n-1}$ with a path on $r+1$ vertices $w_0, w_1, \ldots, w_r$. (See the black vertices in Figure 2.) Then, for each edge xy of $\mathcal{D}_\psi$, we replace $E(x, y)$ in Ψ

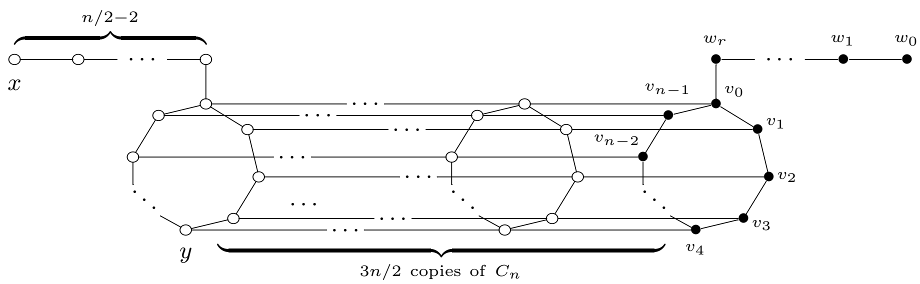

Fig. 2. The gadget for the case of even n where $r = (-n/2 - 2) \bmod (j-1)$.

by the gadget depicted in Figure 2 (consisting of the cartesian product of $\mathcal{C}_n$ and a path on $3n/2$ vertices together with two attached paths on $n/2-2$, resp. $r+1$ vertices). The vertices x and y represent the variables x and y while all other white vertices are new variables, and the black vertices are identified with $v_0, \ldots, v_{n-1}, w_0, \ldots, w_r$ introduced in the previous step.

Finally, we prepend the following quantification to the formula:

$$\exists^{\geq 1} w_0 \, \exists^{\geq j} v_{j-r-2} \, \exists^{\geq j} v_{2j-r-3} \ldots \exists^{\geq j} v_{(k \cdot j - r - k - 1)} \ldots \exists^{\geq j} v_{n/2+1}$$

followed by $\exists^{\geq 1}$ quantification of all the remaining variables of the gadgets.

We prove that $\mathcal{K}_{n/2} \models \Psi$ if and only if $\mathcal{C}_n \models \Psi'$. First, we show that Ψ' correctly simulates the universal quantification of Ψ. The argument for this is essentially the same as in the case of odd n. Next, we need to analyse possible assignments to the vertices $v_0, \ldots, v_{n-1}$. There are two possibilities: either the values chosen for $v_0, \ldots, v_{n-1}$ are all distinct, or not. In the former case, we show that Prover can complete the homomorphism to $\mathcal{C}_n$ if and only $\mathcal{K}_{n/2} \models \Psi$. All other cases are degenerate and we address them separately.

(We skip further details for lack of space; please consult the appendix.) $\square$

Proof of Theorem 2. The case (i) is proved as Proposition 4, and the case (ii) follows from [16]. Finally, the case (iii) is proved as Proposition 5. $\square$

5 Extensions of the CSP

In this section we consider single-quantifier extensions of the classical CSP($\mathcal{B}$), i.e., the evaluation of X-pp sentences, where $X := \{1, j\}$ for some $1 < j \leq |B|$.

5.1 Bipartite graphs

In the case of (irreflexive, undirected) graphs, it is known that $\{1\}$-CSP($\mathcal{H}$) = CSP($\mathcal{H}$) is in L if $\mathcal{H}$ is bipartite and is NP-complete otherwise [16] (for membership in L, one needs also [24]). It is also known that something similar holds for $\{1, |H|\}$-CSP($\mathcal{H}$) = QCSP($\mathcal{H}$) – this problem is in L if $\mathcal{H}$ is bipartite and is NP-hard otherwise [21]. Of course, the fact that $\{1, j\}$-CSP($\mathcal{H}$) is hard on non-bipartite $\mathcal{H}$ is clear, but we will see that it is not always easy on bipartite $\mathcal{H}$.

First, we look at complete bipartite graphs (in a more general statement).

Proposition 6. *Let $\mathcal{K}_{k,l}$ be the complete bipartite graph with partite sets of size k and l. Then $\{1, \ldots, k+l\}$-CSP($\mathcal{K}_{k,l}$) is in* L.

Proof. We reduce to QCSP($\mathcal{K}_2^1$), where $\mathcal{K}_2^1$ indicates $\mathcal{K}_2$ with one vertex named by a constant, say 1. QCSP($\mathcal{K}_2^1$) is equivalent to QCSP($\mathcal{K}_2$) (identify instances if 1 to a single vertex) and both are well-known to be in L (see, e.g., [21]). Let Ψ be input to $\{1,\ldots,k+l\}$-CSP($\mathcal{K}_{k,l}$). Produce Ψ' by substituting quantifiers $\exists^{\geq j}$ with $\exists$, if $j \leq \min\{k,l\}$, or with $\forall$, if $j > \max\{k,l\}$. Variables quantified by $\exists^{\geq j}$ for $\min\{k,l\} < j \leq \max\{k,l\}$ should be replaced by the constant 1. It is easy to see that $\mathcal{K}_{k,l} \models \Psi$ iff $\mathcal{K}_2 \models \Psi'$, and the result follows. $\square$

Proposition 7. *For each j, there exists m s.t. $[2^m 1^*]$-CSP($\mathcal{C}_{2j}$) is* NP*-complete.*

Proof. Membership in NP follows because m is bounded – one may try all possible evaluations to the $\exists^{\geq 2}$ variables. NP-hardness follows as in the proof of case (*iii*) of Theorem 2, but we are reducing from CSP($\mathcal{K}_{n/2}$) not QCSP($\mathcal{K}_{n/2}$). As a consequence, the only instances of $\exists^{\geq 2}$ we need to consider are those used to isolate the cycle $\mathcal{C}_{2j}$ (one may take $m := j+3$). $\square$

Corollary 2. *If $\mathcal{H}$ is bipartite and (for $j \geq 3$) contains some $\mathcal{C}_{2j}$ but no smaller cycle, then exists m s.t. $[2^m 1^*]$-CSP($\mathcal{H}$) is* NP*-complete.*

Proof. Membership and reductions for hardness follow similarly to Proposition 7. The key part is in isolating a copy of the cycle, but we can not do this as easily as before. If d is the diameter of $\mathcal{H}$ (the maximum of the minimal distances between two vertices) then we begin the sentence Ψ' of the reduction with $\exists^{\geq 2} v_1, \ldots, v_{d+1}$, and then, for each $i \in \{1, \ldots, d-j+1\}$ we add $\exists^{\geq 2} x_i, x_i', \ldots, x'^{\cdots\prime}$ ($j-1$ dashes) and join $v_i, \ldots, v_{i+j}, x_i'^{\cdots\prime}, \ldots, x_i$ in a $2i$-cycle (with $E(x_i, v_i)$ also). For each of these $d-j+1$ cycles $\mathcal{C}_{2j}$ we build a separate copy of the rest of the reduction. We can not be sure which of these cycles is evaluated truly on some $\mathcal{C}_{2j}$, but at least one of them must be. $\square$

(See the appendix for an example of why the above construction is necessary.)

In passing, we note the following simple propositions.

Proposition 8. *If $j \in \{2,...,n-3\}$ then one may exhibit a bipartite $\mathcal{H}_j$ of size n such that $\{1,j\}$-CSP($\mathcal{H}_j$) is* Pspace*-complete.*

Proof. The case $j=2$ follows from Theorem 2; assume $j \geq 3$. Take the graph $\mathcal{C}_6$ and construct $\mathcal{H}_j$ as follows. Augment $\mathcal{C}_6$ with $j-3$ independent vertices each with an edge to vertices 1, 3 and 5 of $\mathcal{C}_6$. Apply the proof of Theorem 2 with $\mathcal{H}_j$. $\square$

Proposition 9. *Let $\mathcal{H}$ be bipartite with largest partition in a connected component of size $< j$. Then $\{1,j\}$-CSP($\mathcal{H}$) is in* L*.*

Proof. We will consider an input Ψ to $\{1,j\}$-CSP($\mathcal{H}$) of the form $Q_1 x_1\ Q_2 x_2 \ldots Q_m x_m\ \psi(x_1, x_2, \ldots, x_m)$. An instance of an $\exists^{\geq j}$ variable is called *trivial* if it has neither a path to another (distinct) $\exists^{\geq j}$ variable, nor a path to an $\exists$ variable that precedes it in the natural order on $\mathcal{D}_\psi$. The key observation here is that any non-trivial $\exists^{\geq j}$ variable *must* be evaluated on more than one partition of a connected component. If in Ψ there is a non-trivial $\exists^{\geq j}$ variable, then Ψ must be a no-instance (as $\exists^{\geq j}$s must be evaluated on more than one partition of a connected component, and a path can not be both even and odd in length). All other instances are readily seen to be satisfiable. Detecting if Ψ contains a non-trivial $\exists^{\geq j}$ variable is in L by [24], and the result follows. $\square$

We note that Proposition 8 is tight, namely in that $\{1, j\}$-CSP($\mathcal{H}$) is in L if $j \in \{1, |H|-2, |H|-1, |H|\}$. (For the proof, please consult the appendix.)

Proposition 10. *If $\mathcal{H}$ is bipartite and contains $\mathcal{C}_4$, then $\Psi \in \{1,2\}$-CSP($\mathcal{C}_4$) iff the underlying graph $\mathcal{D}_\psi$ of Ψ is bipartite. In particular, $\{1,2\}$-CSP($\mathcal{H}$) is in L.*

Proof. Necessity is clear; sufficiency follows by the canonical evaluation of $\exists^{\geq 1}$ and $\exists^{\geq 2}$ on a fixed copy of $\mathcal{C}_4$ in $\mathcal{H}$. Membership in L follows from [24]. □

Proposition 11. *Let $\mathcal{H}$ be a forest, then $[2^m 1^*]$-CSP($\mathcal{H}$) is in* P *for all m.*

Proof. We evaluate each of the m variables bound by $\exists^{\geq 2}$ to all possible pairs, and what we obtain in each case is an instance of CSP($\mathcal{H}'$) where $\mathcal{H}'$ is an expansion of $\mathcal{H}$ by some constants, i.e., equivalent to the retraction problem. It is known that Ret($\mathcal{H}$) is in P for all forests $\mathcal{H}$ [14], and the result follows. □

We bring together some previous results into a classification theorem.

Theorem 3. *Let $\mathcal{H}$ be a graph. Then*
- $[2^m 1^*]$-*CSP($\mathcal{H}$) $\in$* P *for all m, if $\mathcal{H}$ is a forest or a bipartite graph containing $\mathcal{C}_4$*
- $[2^m 1^*]$-*CSP($\mathcal{H}$) is NP-complete from some m, if otherwise.*

Proof. Membership of NP follows since m is fixed. The cases in P follow from Propositions 11 and 10. Hardness for non-bipartite graphs follows from [16] and for the remaining bipartite graphs it follows from Corollary 2. □

6 The complexity of QCSP($\mathcal{C}_4^*$)

Let $\mathcal{C}_4^*$ be the reflexive 4-cycle. The complexities of Ret($\mathcal{C}_6$) and Ret($\mathcal{C}_4^*$) are both hard (NP-complete) [13,12], and retraction is recognised to be a "cousin" of QCSP (see [2]). The problem QCSP($\mathcal{C}_6$) is known to be in L (see [21]), but the complexity of QCSP($\mathcal{C}_4^*$) was hitherto unknown. Perhaps surprisingly, we show that is is markedly different from that of QCSP($\mathcal{C}_6$), being Pspace-complete.

Proposition 12. *$\{1,2,3,4\}$-CSP($\mathcal{C}_4^*$) is* Pspace-*complete.*

Corollary 3. *QCSP($\mathcal{C}_4^*$) is* Pspace-*complete.*

The proofs of these claims are based on the hardness of the retraction problem to reflexive cycles [12] and are similar to our proof of the even case of Proposition 5. (We skip the details for lack of space; please consult the appendix.)

While QCSP($\mathcal{C}_4^*$) has different complexity from QCSP($\mathcal{C}_6$), we remark that the better analog of the retraction complexities is perhaps that $\{1, |C_4^*|\}$-CSP($\mathcal{C}_4^*$) and $\{1, |C_6|/2\}$-CSP($\mathcal{C}_6$) **do** have the same complexities (recall the reductions to Ret($\mathcal{C}_4^*$) and Ret($\mathcal{C}_6$) involved CSP($\mathcal{K}_{|C_4^*|}$) and CSP($\mathcal{K}_{|C_6|/2}$), respectively.

7 Conclusion

We have taken first important steps to understanding the complexity of CSPs with counting quantifiers, even though several interesting questions have resisted solution. We would like to close the paper with some open problems.

In Section 4.1, the case $n = 2j$ remains. When $j = 1$ and $n = 2$, we have $\{1\}$-CSP($\mathcal{K}_2$)=CSP($\mathcal{K}_2$) which is in L by [24]. For higher j, the question of the complexity of $\{j\}$-CSP($\mathcal{K}_{2j}$) is both challenging and open.

We would like to prove the following more natural variants of Theorem 3, whose involved combinatorics appear to be much harder.

Conjecture 1. Let $\mathcal{H}$ be a graph. Then

– $[2^*1^*]$-CSP($\mathcal{H}$) is in P, if $\mathcal{H}$ is a forest or a bipartite graph containing $\mathcal{C}_4$,
– $[2^*1^*]$-CSP($\mathcal{H}$) is NP-hard, if otherwise.

Conjecture 2. Let $\mathcal{H}$ be a graph. Then

– $\{1, 2\}$-CSP($\mathcal{H}$) is in P, if $\mathcal{H}$ is a forest or a bipartite graph containing $\mathcal{C}_4$,
– $\{1, 2\}$-CSP($\mathcal{H}$) is NP-hard, if otherwise.

References

1. Barto, L., Kozik, M., and Niven, T. The CSP dichotomy holds for digraphs with no sources and no sinks (a positive answer to a conjecture of Bang-Jensen and Hell). *SIAM Journal on Computing 38*, 5 (2009), 1782–1802.
2. Bodirsky, M., Kára, J., and Martin, B. The complexity of surjective homomorphism problems – a survey. *CoRR abs/1104.5257* (2011).
3. Börner, F., Bulatov, A. A., Chen, H., Jeavons, P., and Krokhin, A. A. The complexity of constraint satisfaction games and qcsp. *Inf. Comput. 207*, 9 (2009), 923–944.
4. Börner, F., Krokhin, A., Bulatov, A., and Jeavons, P. Quantified constraints and surjective polymorphisms. Tech. Rep. PRG-RR-02-11, Oxford University, 2002.
5. Bulatov, A. A dichotomy theorem for constraint satisfaction problems on a 3-element set. *J. ACM 53*, 1 (2006), 66–120.
6. Bulatov, A., Krokhin, A., and Jeavons, P. G. Classifying the complexity of constraints using finite algebras. *SIAM Journal on Computing 34* (2005), 720–742.
7. Chen, H. The complexity of quantified constraint satisfaction: Collapsibility, sink algebras, and the three-element case. *SIAM J. Comput. 37*, 5 (2008), 1674–1701.
8. Chen, H. Existentially restricted quantified constraint satisfaction. *Inf. Comput. 207*, 3 (2009), 369–388.
9. Chen, H. Quantified Constraint Satisfaction and the Polynomially Generated Powers Property. *Algebra Universalis*, 65 (2011), 213–241
10. Clark, R., and Grossman, M. Number sense and quantifier interpretation. *Topoi 26*, 1 (2007), 51–62.
11. Ebbinghaus, H.-D., and Flum, J. *Finite Model Theory*. Springer, 1999. 2nd edition.
12. Feder, T., and Hell, P. List homomorphisms to reflexive graphs. *J. Comb. Theory, Ser. B 72*, 2 (1998), 236–250.
13. Feder, T., Hell, P., and Huang, J. List homomorphisms and circular arc graphs. *Combinatorica 19*, 4 (1999), 487–505.
14. Feder, T., Hell, P., Jonsson, P., Krokhin, A. A., and Nordh, G. Retractions to pseudoforests. *SIAM J. Discrete Math. 24*, 1 (2010), 101–112.
15. Feder, T., and Vardi, M. The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory. *SIAM Journal on Computing 28* (1999), 57–104.
16. Hell, P., and Nešetřil, J. On the complexity of H-coloring. *Journal of Combinatorial Theory, Series B 48* (1990), 92–110.
17. Hemaspaandra, E. Dichotomy theorems for alternation-bounded quantified boolean formulas. *CoRR cs.CC/0406006* (2004).
18. Kolaitis, P. G., and Vardi, M. Y. *Finite Model Theory and Its Applications (Texts in Theoretical Computer Science. An EATCS Series)*. Springer-Verlag New York, Inc., 2005, ch. A logical Approach to Constraint Satisfaction.
19. Lange, K.-J. Some results on majority quantifiers over words. In *Computational Complexity, 2004. Proceedings. 19th IEEE Annual Conference on* (june 2004), pp. 123 – 129.
20. Martin, B. First order model checking problems parameterized by the model. In *CiE 2008, LNCS 5028* (2008), pp. 417–427.
21. Martin, B., and Madelaine, F. Towards a trichotomy for quantified H-coloring. In *2nd Conf. on Computatibility in Europe, LNCS 3988* (2006), pp. 342–352.
22. Otto, M. *Bounded variable logics and counting – A study in finite models*, vol. 9. Springer-Verlag, 1997. IX+183 pages.
23. Papadimitriou, C. H. *Computational Complexity*. Addison-Wesley, 1994.
24. Reingold, O. Undirected connectivity in log-space. *J. ACM 55*, 4 (2008), 1–24.
25. Schaefer, T. J. The complexity of satisfiability problems. In *Proceedings of STOC'78* (1978), pp. 216–226.

A Appendix

A.1 Omitted proofs

Proof of Proposition 1. We show that $\mathcal{K}_{\binom{2j+1}{j}} \models \Psi$ iff $\mathcal{K}_{2j+1} \models \Psi'$. Observe there is a natural bijection π from subsets of j elements of $\mathcal{K}_n$ to vertices of $\mathcal{K}_{\binom{2j+1}{j}}$. In the simulation of QCSP($\mathcal{K}_{\binom{2j+1}{j}}$) in $\{j\}$-CSP($\mathcal{K}_{2j+1}$), Adversary may be seen to take on the role of denying $\mathcal{K}_{\binom{2j+1}{j}} \models \Psi$ while Prover is asserting that it is true. Thus, Adversary may always be assumed to play variables $v^1, \ldots, v^j$ such that $|\{v^1, \ldots, v^j\}| = j$, because otherwise he is simply making the job of Prover easier (by the properties of the gadget $\mathcal{G}_j$). The behaviour of existential quantification in the simulation is easy to see, but we will consider more carefully the behaviour of universal quantification. The additional $v^{1,1}, \ldots, v^{1,j+1}, \ldots, v^{j,1}, \ldots, v^{j,j+1}$ force that every possible subset $\{a_1, \ldots, a_j\} \subset \{1, \ldots, 2j+1\}$ can be forced by Adversary on $v^1, \ldots, v^j$. Indeed, Adversary may force any single element on v^i by avoiding it in $v^{i,1}, \ldots, v^{i,j+1}$.

($\Rightarrow$) Assume $\mathcal{K}_{\binom{2j+1}{j}} \models \Psi$. However Prover plays the variables in Ψ' corresponding to universal variables of Ψ, she will be able to find a witness set $\pi^{-1}(a)$ for the variables in Ψ' corresponding to an existential variable x in Ψ, precisely because that existential variable has some witness $a \in K_{\binom{2j+1}{j}}$.

($\Leftarrow$) Assume $\mathcal{K}_{2j+1} \models \Psi'$. No matter how Prover plays to win $\mathscr{G}(\Psi', \mathcal{K}_{2j+1})$, she will have possible witnesses sets $\{a_1, \ldots, a_j\}$ for variables $\{v^1, \ldots, v^j\}$ in Ψ' corresponding to an existential variable v of Ψ, for all sets $\{b_1, \ldots, b_j\} \subset \{1, \ldots, 2j+1\}$ corresponding to universal variables $\{u_1, \ldots, v_j\}$ of Ψ (because of the behaviour of the universal variable simulation). Thus the existential witness $\pi(\{a_1, \ldots, a_j\})$ may be used in Ψ for v, and the result follows. $\square$

Proof of the claims (1a-c) from Proposition 4.

For (1a), let y be a predecessor of x. Then for the value i chosen by Adversary for y, Prover must offer a set of at least three vertices of $\mathcal{C}_n$ that are adjacent to i in $\mathcal{C}_n$. Since there are only two such vertices, Adversary can always choose for x a vertex non-adjacent to i at which point Prover loses.

For (1b), let y be the first vertex of the connected component of $\mathcal{D}_\psi$ that contains x. Assume $y \neq x$ and consider the path $\mathcal{P}$ between x and y in $\mathcal{D}_\psi$. Without loss of generality, assume that the value i chosen by Adversary for y is even. Note that, because n is even, if the length of $\mathcal{P}$ is also even, then Adversary must choose an even value for x, while if the length is odd, she must choose an odd value (otherwise Prover loses). However, as $j > n/2$, the set provided by Prover for x contains both an even and an odd number. Thus Adversary is allowed to choose for x the wrong parity and Prover loses.

For (1c), suppose that y and z with $y \prec z$ are predecessors of x where z is quantified by $\exists^{\geq j'}$ for some $j' \geq 2$. If i is the value chosen by Adversary for y, then Prover must offer for z a set of $j' \geq 2$ values which hence must contain at least one value different from i. Adversary then chooses this value i' after which Prover must offer for x two distinct vertices i'', i''' of $\mathcal{C}_n$ adjacent to both i and i'. But then i, i', i'', i'' yield a 4-cycle in $\mathcal{C}_n$, impossible if $n \neq 4$. $\square$

Proof of the case of odd n in Proposition 5. We argue that $\mathcal{C}_n \models \Psi$ if and only if $\mathcal{C}_n \models \Psi'$. To do this, it suffices to show that Ψ' correctly simulates the universal quantifiers of Ψ. Namely, it suffices to prove that $\mathcal{C}_n \models Q_x\pi_x$, and for each $\ell \in \{0 \dots n-1\}$, Adversary has a strategy on $Q_x\pi_x$ that evaluates x to ℓ.

For the first part, we provide a strategy for Prover. We treat x as x_1^{n+1}. For x_1^1, Prover offers any set. For x_1^k where $k \geq 2$, let i be the value chosen by Adversary for x_1^{k-1}. By Lemma 2, we observe that there are exactly j vertices in $\mathcal{C}_n$ having a walk to i of length $j-1$. Prover offers this set for x_1^k. This allows her to choose values for $x_2^{k-1} \dots x_{j-1}^{k-1}$ as the path $x_1^{k-1}, \dots x_{j-1}^{k-1}, x_1^k$ encodes precisely the fact that there exists a walk of length $j-1$ between the values chosen for x_1^{k-1} and x_1^k. Thus $\mathcal{C}_n \models Q_x\pi_x$.

For the second part, consider any $\ell \in \{0 \dots n-1\}$. We explain a strategy for Adversary that allows him to choose ℓ for x. First, Adversary chooses any value for x_1^1. Let i_0 be this value, and by the second part of Lemma 2, choose a sequence of n numbers $i_1, i_2, \dots, i_n$ either all from $\{-1,+1\}$ if j is odd, or all from $\{-2,0,+2\}$ if j is even, such that $i_0 + i_1 + i_2 + \dots + i_n = \ell$. After this, consider inductively $k \geq 2$ and let i be the value chosen by Adversary for x_1^{k-1}. By Lemma 2, there are exactly j possible values that Prover can offer if she does not want to lose. Thus Prover is forced to offer all these values. In particular, if j is even, this set contains values $i+1$ and $i-1$ (mod n) while if j is odd, the set contains values $i+2$, i, and $i-2$ (mod n). Thus Adversary is allowed to choose the value $i + i_{k-1}$ (mod n) for x_1^k. This shows that Adversary is allowed to choose the value $i_0 + i_1 + \dots + i_n = \ell$ for $x_1^{n+1} = x$.

Thus, this proves that Ψ' correctly simulates the universal quantifiers of Ψ, and consequently $\mathcal{C}_n \models \Psi$ if and only if $\mathcal{C}_n \models \Psi'$. For odd n, this completes the proof of the claim that $\{1,j\}$-CSP($\mathcal{C}_n$) is Pspace-hard. $\square$

Proof of the case of even n in Proposition 5.

We analyse possible assignments to the vertices $v_0, \dots, v_{n-1}$. For clarity, we define $\alpha_k = kj - r - k - 1$ and note that $n/2 + 1 = \alpha_k$ for $k = \lceil \frac{n+4}{2(j-1)} \rceil$. By the symmetry of $\mathcal{C}_n$, we assume that Adversary chooses for w_0 the value $n - r - 1$. The next quantified vertex is $v_{j-r-2} = v_{\alpha_1}$ in distance $j-1$ from w_0. Thus, by Lemma 2, there are exactly j values that Prover can and must offer. Among them, we find $j - r - 2 = \alpha_1$. Similarly, for $2 \leq k \leq \lceil \frac{n+4}{2(j-1)} \rceil$, the vertex $v_{\alpha_{k-1}}$ is in distance $j-1$ from v_{α_k}, and hence, Prover is forced to offer a set of j values only depending on the value chosen for $v_{\alpha_{k-1}}$. In particular, if α_{k-1} was chosen for $v_{\alpha_{k-1}}$, then Adversary can choose α_k for v_{α_k}. This argument also shows that if Prover acts as we describe, then in every possible case she can complete the homomorphism for the path $w_0, w_1, \dots, w_r, v_0 \dots, v_{n/2+1}$. Further, she also has a way of assigning the values to $v_{n/2+2}, \dots, v_{n-1}$. This can be seen as follows. First, note that the distance betwen $v_{n/2+1}$ and v_0 is $n/2 - 1$. Thus, if $n/2$ is odd, then the values assigned to v_0 and $v_{n/2+1}$ have the same parity because $n/2 + 1$ is even and we observe that between any two vertices of the same parity in $\mathcal{C}_n$ there exists a walk of length $n/2 - 1$. Similarly, if $n/2$ is even, the values chosen for $v_0, v_{n/2+1}$ have different parity and between any two vertices of $\mathcal{C}_n$ of different parity there is a walk of length $n/2 - 1$.

This concludes the argument for the vertices $v_0, \ldots, v_{n-1}$. It implies two possible types of outcomes: either the values chosen for $v_0, \ldots, v_{n-1}$ are all distinct, or not. To obtain the former case, for each $1 \leq k \leq \lceil \frac{n+4}{2(j-1)} \rceil$, Adversary chooses α_k for v_{α_k}. This forces assigning i to v_i for all $i \in \{0, \ldots, n/2+1\}$ and thus consequently also for all the other v_i's. We shall assume this situation first. For all other (degenerate) cases we use a different argument explained later.

Thus assuming that $v_0, \ldots, v_{n-1}$ get assigned values $0, \ldots, n-1$ in that order and the values for the original variables of Ψ are chosen, we argue that Prover can finish the homomorphism if and only if the assignment to the variables of Ψ is a proper colouring for $\mathcal{D}_\psi$. This follows exactly as in [13]. Namely, in every gadget, each copy of $\mathcal{C}_n$ is forced to copy the assignment from the adjacent copy of $\mathcal{C}_n$, shifted by $+1$ or by -1 (mod n). In particular, if i is the value assigned to y, the vertex z opposite y in the last copy of $\mathcal{C}_n$ is necessarily assigned value $n/2+i$ (mod n). This implies that the value assigned to x is different from i as the path from z to x is too short (of length less than $n/2$). On the other hand, this path is long enough so that any value of the same parity as i but different from i can be chosen for x such that the homomorphism can be completed. This precisely simulates the edge predicate of Ψ. Finally, we observe that Prover can choose whether consecutive copies of $\mathcal{C}_n$ are shifted by $+1$ or -1 and there are exactly $3n/2$ copies of $\mathcal{C}_n$. Thus, by Lemma 2, every possible odd number from $\{0, \ldots, n-1\}$ can be chosen for y by a particular series of shifts. It follows that $\{1, 3, \ldots, n-1\}$ is precisely the set colours we use to simulate QCSP$(\mathcal{K}_{n/2})$.

Now, we discuss the degenerate cases. Namely, we show that, regardless of the assignment to $v_0, \ldots, v_{n-1}$, for each copy of the gadget (in Figure 2) there is a way to complete the homomorphism (by assigning the values to the white vertices) in such a way that if ℓ is the value assigned to y, then the vertex opposite y in the last copy of C_n is assigned value $\ell + n/2$ (mod n). As this is exactly what happens in the non-degenerate case, the rest will follow. Note that consideration of these degenerate cases is the reason we use a chain of $3n/2$ copies of $\mathcal{C}_n$ in the gadget of Figure 2, instead of the $n/2$ used in the like gadget in [12].

Since we assume that the vertices $v_0, \ldots, v_{n-1}$ are assigned a proper subset of $\{0, \ldots, n-1\}$, it can be seen that they are assigned a circularly consecutive subset of these numbers, and this subset is of size at most $n/2+1$ as otherwise the assignment cannot be a homomorphism. (Recall that in the proof we argue that we can assume that the assignment to $v_0, \ldots, v_{n-1}$ is a homomorphism).

For simplicity, let λ denote the assignment constructed so far, i.e., a mapping from the assigned vertices to their assigned values. We explain how to complete this assignment for the gadget so that it becomes a homomorphism to C_n.

The gadget contains $3n/2$ copies of C_n. We consider them from the right to left, namely $\{v_0, \ldots, v_{n-1}\}$ is the 1st copy, and the $3n/2$-th copy is the one containing y. With this in mind, we denote by $v_0^i, \ldots, v_{n-1}^i$ the respective copies of $v_0, \ldots, v_{n-1}$ in the i-th copy of C_n. In particular, y is the vertex $v_{n/2}^{3n/2}$.

We describe the assignment to the copies of C_n in three phases. In the first phase, we assign values to the first $n/2$ copies. Consider $1 \leq i < n/2$, and assume

that the vertices $v_0^i, \ldots, v_{n-1}^i$ are assigned values between a and b (inclusive) in the clock-wise order. Then the assignment to the $(i+1)$-st copy of C_n is as follows. For $k \in \{0, \ldots, n-1\}$, if $\lambda(v_k^i) = a$, then we set $\lambda(v_k^{i+1}) = a + 1 \pmod n$, otherwise we set $\lambda(v_k^{i+1}) = \lambda(v_k^i) - 1 \pmod n$. It is easy to verify that this constitutes a homomorphism to C_n. It follows that $\left|\lambda\left(\{v_0^{n/2}, \ldots, v_{n-1}^{n/2}\}\right)\right| = 2$.

Next, we explain the assignment to the second $n/2$ copies of C_n. Let ℓ be the value assigned to y. We choose the values for the second $n/2$ copies in such a way that consecutive copies of C_n are just shifted by $+1$ or -1. We can choose an appropriate sequence of $+1, -1$ shifts so that the value assigned to $v_{n/2}^n$ is exactly $\ell + n/2 \pmod n$. (The argument about the parity of these values is the same as in the non-degenerate case.) We further conclude that $\left|\lambda\left(\{v_0^n, \ldots, v_{n-1}^n\}\right)\right| = 2$.

The assignment to the final $n/2$ copies is as follows. For $n \le i < 3n/2$, again assume that the vertices $v_0^i, \ldots, v_{n-1}^i$ are assigned values between a and b in the clockwise order. Then for $k \in \{0, \ldots, n-1\} \setminus \{n/2\}$, if $\lambda(v_k^i) = a+1 \pmod n$ and $\lambda\left(v^i_{(k-1) \bmod n}\right) = \lambda\left(v^i_{(k+1) \bmod n}\right) = a$, then we set $\lambda(v_k^{i+1}) = a$, and otherwise we set $\lambda(v_k^{i+1}) = \lambda(v_k^i) + 1$. Again, we conclude that this constitutes a homomorphism, and it follows that $\left|\lambda\left(\{v_0^{3n/2}, \ldots, v_{n-1}^{3n/2}\}\right)\right| = n/2 + 1$. In particular, we observe that $\lambda\left(y = v_{n/2}^{3n/2}\right) = \ell$ and $\lambda\left(v_0^{3n/2}\right) = \ell + n/2 \pmod n$.

That concludes the argument. □

Proof of Proposition 12. We will reduce from the problem QCSP($\mathcal{K}_4$) (known to be Pspace-complete from, e.g., [3]). We will borrow heavily from the reduction of CSP($\mathcal{K}_4$) to Ret($\mathcal{C}_4^*$) in [12]. The reduction has a very similar flavour to that used in Case (*ii*) of Theorem 2, but borrows from [12] instead of [13].

For an input $\Psi := Q_1 x_1 Q_2 x_2 \ldots Q_m x_m \; \psi(x_1, x_2, \ldots, x_m)$ for QCSP($\mathcal{K}_4$) we build an input Ψ' for $\{1,2,3,4\}$-CSP($\mathcal{C}_4^*$) as follows. We begin by considering the graph $\mathcal{D}_\psi$, from which we first build a graph $\mathcal{G}' := \mathcal{D}_\psi \uplus \mathcal{C}_4^*$. Now we build $\mathcal{G}''$ from $\mathcal{G}'$ by replacing every edge $(x, y) \in \mathcal{D}_\psi$ with the following gadget (which connects also to the fixed copy of $\mathcal{C}_4^*$ in $\mathcal{G}'$ – induced by $\{z_1, \ldots, z_4\}$ – as drawn in the picture).

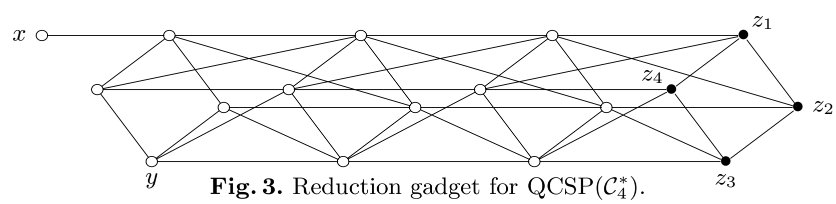

Fig. 3. Reduction gadget for QCSP($\mathcal{C}_4^*$).

$\phi_{\mathcal{G}''}$ will form the quantifier-free part of Ψ'; we now explain the structure of the quantifiers. Let $z_1, \ldots, z_4$ correspond in $\phi_{\mathcal{G}''}$ to the fixed copy of $\mathcal{C}_4^*$ in $\mathcal{G}''$. Ψ' begins $\exists z_1 \exists^{\ge 2} z_2 \exists^{\ge 3} z_3 \exists^{\ge 2} z_4$ (z_1 could equally be quantified with $\exists^{\ge j}$, $j > 1$). Now we continue in the quantifier order of Ψ. When we meet an $\exists$ quantifier, we quantify with $\exists$ the corresponding vertex in $\phi_{\mathcal{G}''}$. When we meet a $\forall$ quantifier, we quantify with $\forall = \exists^{\ge 4}$ the corresponding vertex in $\phi_{\mathcal{G}''}$. Finally, we quantify with $\exists$ all remaining variables, corresponding to vertices we added in gadgets

in $\mathcal{G}''$. We claim that $\mathcal{K}_3 \models \Psi$ iff $\mathcal{C}_4^* \models \Psi'$. The proof of this proceeds as with Theorem 2 (though there are several more degenerate cases to consider). $\square$

Proof of Corollary 3. We give a reduction from $\{1,2,3,4\}$-CSP($\mathcal{C}_4^*$) to QCSP($\mathcal{C}_4^*$), using the following shorthands (x', x'' must appear nowhere else in ϕ, which may contain other free variables).

$$\begin{aligned}
\exists^{\geq 1} x\ \phi(x) &:= \exists x\ \phi(x)\\
\exists^{\geq 2} x\ \phi(x) &:= \forall x' \exists x\ E(x', x) \wedge \phi(x)\\
\exists^{\geq 3} x\ \phi(x) &:= \forall x'' \forall x' \exists x\ E(x'', x) \wedge E(x', x) \wedge \phi(x)\\
\exists^{\geq 4} x\ \phi(x) &:= \forall x\ \phi(x)
\end{aligned}$$

On $\mathcal{C}_4^*$, it is easy to verify that, for each $i \in [4]$, $\exists^i x\ \phi(x)$ holds iff there exist at least i elements x satisfying ϕ. The result follows easily (note that each use of shorthand substitution involves new variables corresponding to x and x' above).

That concludes the proof. $\square$

A.2 Example for Corollary 2

As an example of why the construction of Corollary 2 was necessary, consider the graph $\mathcal{H}_6^{\text{hairy}}$, depicted in Figure 4 on the right, with 18 vertices made from $\mathcal{C}_6$ with twin paths of length one added to each vertex. The first part of Ψ' takes the form depicted in Figure 4 on the left.

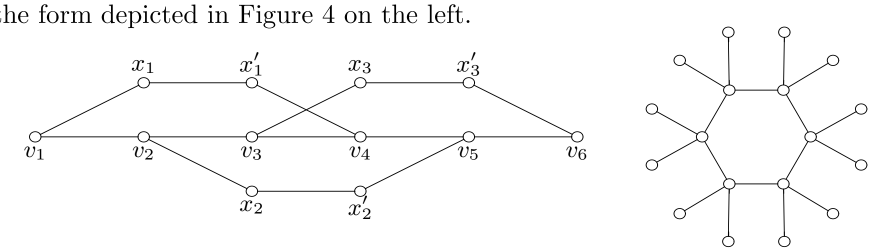

Fig. 4. Finding $\mathcal{C}_6$ in $\mathcal{H}_6^{\text{hairy}}$.

In this case, each of $v_1, v_2, v_3, v_4, x_1', x_1$ or $v_2, v_3, v_4, v_5, x_2', x_2$ must be evaluated to $\mathcal{C}_6$ ($v_3, v_4, v_5, v_6, x_3', x_3$ need not). For an example where $v_1, v_2, v_3, v_4, x_1', x_1$ may fail to be evaluated to a $\mathcal{C}_6$, but a later cycle does not, consider two disjoint copies of $\mathcal{C}_6$ joined by a path of length two.

A.3 Tightness of Proposition 8

Proposition 13. $\{1,3\}$-*CSP*($\mathcal{P}_5$) *is in* L.

Proof. We will consider an input Ψ to $\{1,3\}$-CSP($\mathcal{P}_5$) of the form $Q_1 x_1 Q_2 x_2 \ldots Q_m x_m\ \psi(x_1, x_2, \ldots, x_m)$. An instance of an $\exists^{\geq 3}$ variable is called *trivial* if it has neither a path to another (distinct) $\exists^{\geq 3}$ variable, nor a path to an $\exists$ variable that precedes it in the natural order on $\mathcal{D}_\psi$. A minimum requirement on a yes-instance Ψ is that $\mathcal{D}_\psi$ be bipartite, which can be determined in L; henceforth we

assume this. Note that in the game $\mathscr{G}(\Psi, \mathcal{P}_5)$, for $\exists$ variables, Prover advances a singleton set (and Adversary must choose that vertex); i.e. effectively only Prover plays.

In this case one partition is a vertex larger than the other, and this partition contains the vertices at both ends. Any non-trivial instance of an $\exists^{\geq 3}$ variable must be evaluable on all vertices of the larger partition. If in $\mathcal{D}_\psi$ there are two $\exists^{\geq 3}$ vertices with a path of odd length or length < 4 between them, then Ψ must be a no-instance (recall these variables must be evaluable on all vertices of the larger partition, at extreme ends these are at distance 4). Further, if in $\mathcal{D}_\psi$ there is an $\exists$ vertex preceding in the order an $\exists^{\geq 3}$ vertex, and these are joined by a path of length < 2, then Ψ must be a no-instance. Otherwise, we claim Ψ is a yes-instance, and Prover may witness this by using the following *centre-finding* strategy. Let 3 be the centre vertex of $\mathcal{P}_5$ and let 2 be one of its neighbours. Essentially, Prover is always trying to move towards 3 and 2 on the $\exists$ variables (recall she always suggests the three variables in the larger bipartition as the witness set for an $\exists^{\geq 3}$ variable). When given an $\exists$ variable, corresponding to a vertex x in $\mathcal{D}_\psi$, to evaluate, Prover must look at all vertices at distance ≤ 2 from x in $\mathcal{D}_\psi$. If some of these are already evaluated, then Prover looks at the closest she can get to 3 and 2 with x (bearing parity in mind). This will result in x being played on one of the vertices $2, 3, 4$; and on 4 only if x is adjacent to a vertex already played on 5. Otherwise, if none of these are already played, then Prover looks to see if there is a path in $\mathcal{D}_\psi$ to either an $\exists^{\geq 3}$ vertex, as yet unplayed, or any other vertex already played. If there is such a path to an already played vertex, then she plays on 3 or 2 according to the parity of the played vertex and length of the path. If there is such a path to an unplayed $\exists^{\geq 3}$ vertex, and it is of odd length, Prover plays x on 2. In all other cases, Prover plays x on 3. As this always provides a vertex on which Prover may play, it is seen to be a winning strategy for her. $\square$

Proposition 14. *Let $\mathcal{H}$ be bipartite. Then $\{1, j\}$-CSP($\mathcal{H}$) is in* L *if* $j \in \{1, |H|-2, |H|-1, |H|\}$.

Proof. When $j := |H|$ we have QCSP($\mathcal{H}$) and we refer for the result to [21]. For $j := |H|-1$, we argue as in the proof of Proposition 9 unless $\mathcal{H}$ is the complete bipartite (star) $\mathcal{K}_{1,l}$ (for some l), in which case we appeal to Proposition 6. The case $j := |H|-2$ is not much more complicated. If we do not fall as in the proof of Proposition 9 or under Proposition 6, then we are equivalent to $\{1, 3\}$-CSP($\mathcal{P}_5$) and the result follows from Proposition 13. $\square$